



Efficient and Rate-Distortion Optimal Wavelet Packet Basis Selection in JPEG2000

Thomas Stutz, Andreas Uhl

► To cite this version:

Thomas Stutz, Andreas Uhl. Efficient and Rate-Distortion Optimal Wavelet Packet Basis Selection in JPEG2000. IEEE Transactions on Multimedia, 2012, 14 (2), pp.264-277. 10.1109/TMM.2011.2177644 . hal-00669367

HAL Id: hal-00669367

<https://hal.science/hal-00669367>

Submitted on 13 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficient and Rate-Distortion Optimal Wavelet Packet Basis Selection in JPEG2000

Thomas Stütz and Andreas Uhl

Abstract—This paper discusses optimal wavelet packet basis selection within JPEG2000. Algorithms for rate-distortion optimal wavelet packet basis selection in JPEG2000 are presented and compared to more efficient wavelet packet basis selection schemes. Both isotropic and anisotropic wavelet packet bases are considered. For the first time, computationally efficient heuristics are compared to the best bases in the standardized coding framework of JPEG2000. For the first time, the maximum performance gains of custom wavelet packets in JPEG2000 can be assessed. The algorithms are evaluated on a wide range of highly textured image data.

Index Terms—Image compression, JPEG2000, rate-distortion optimization, wavelet packet bases.

I. INTRODUCTION

WAVELET packet bases (WPBs) [1] offer to adapt the wavelet transform to the source signal (image) characteristics and thus improve the compression performance. WPBs are an alternative to the classical dyadic wavelet decomposition (also referred to as pyramidal) and allow to further decompose all subbands and not just the LL subband, which leads to an enormous number of possible WPBs. The application of an adapted wavelet packet basis (WPB) for image compression purposes has been subject to investigation since the introduction of the first feasible selection technique called “best basis algorithm” [1]. A brute-force search for the best WPB is computationally infeasible even for moderate maximum decomposition depths; for 2-D signals and wavelet decomposition depth 5 there are 5.6×10^{78} possible isotropic WPBs. In Fig. 1 examples of WPBs (compare to the original images in Fig. 2) are shown; in Fig. 1(c) the best anisotropic WPB for the artificial image is illustrated.

The approach of [1] employs a rate-independent but suboptimal basis selection scheme, which is based on various additive cost functions which only estimate the actual coding cost. An extension to this approach employing non-additive cost functions has been developed soon after [2]. Genetic algorithms have been used [3] to assess the degree of optimality and to further

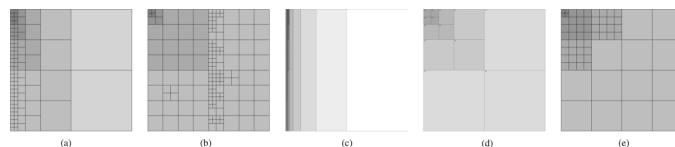


Fig. 1. Best WPBs for specific images and the WSQ-WPB. (a) Artificial. (b) Barbara. (c) Aniso Artificial. (d) Lena. (e) WSQ.

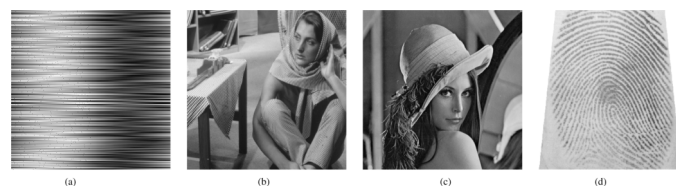


Fig. 2. Selected test images. (a) Artificial. (b) Barbara. (c) Lena. (d) Fingerprint.

optimize the subband structures found by the best basis algorithms proposed in earlier work.

The employment of rate-distortion optimization criteria for WPB selection has been first demonstrated for classical wavelet-based compression schemes [4]. For certain compression schemes, a certain source image, and a specific target bitrate, the optimal WPB can be computed in feasible time. For zero-tree-based compression algorithms, a Markov chain-based cost function estimating the cost of zero tree coding has been employed to find well suited WPBs [5]. In previous work [6], a proprietary wavelet block-based compression scheme has been introduced incorporating the principle of [4] for WPB selection. Subsequent works of the authors [7], [8] propose fast and efficient basis selection methods for their proprietary compression system with a lower computational complexity connected with a little loss of rate-distortion performance in comparison with the original work.

In recent work [9], isotropic and anisotropic wavelet packet decompositions of the dual-tree wavelet transform have been successfully applied for image coding. The approach [9] outperforms JPEG2000 Part 1 (it employs a different entropy coder). Compared to the approach presented in this paper, a main difference is the application of a different transform in [9], that although redundant delivers efficient/sparse representations of the input signal, that can even outperform the compaction properties of the best wavelet packets of the classic DWT (discrete wavelet transform). The approach presented in this paper, however, has the great advantage that it can be implemented and deployed within an already standardized compression framework, namely JPEG2000 Part 2.

Manuscript received October 31, 2010; revised April 03, 2011 and July 29, 2011; accepted October 04, 2011. Date of publication December 06, 2011; date of current version March 21, 2012. This work was supported by the Austrian Science Fund under Project 19159. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Oscar C. Au.

T. Stütz is with the IRCCyN UMR CNRS 6597, Polytech'Nantes LUNAM Université, Université de Nantes, Nantes, France (e-mail: thomas.w.stuetz@gmail.com).

A. Uhl is with the Department of Computer Sciences, University of Salzburg, Salzburg, Austria (e-mail: uhl@cosy.sbg.ac.at).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2011.2177644

The main application field of WPBs in image compression are textured data, with many contributions devoted to fingerprint images. Fingerprint images exhibit characteristically high energy in certain high frequency bands resulting from the ridge-valley patterns. To account for this property, the WSQ standard for lossy fingerprint compression as adopted by the FBI [10], [11] uses a specific wavelet packet subband structure which emphasizes high frequency bands. Inspired by this algorithm, a few WP-based fingerprint compression schemes have been developed (e.g., [12]–[14]).

JPEG2000 Part 1 is currently gaining more and more market acceptance (DCI, DICOM, digital libraries), research on the actual capabilities of the extensions of Part 2 are of considerable practical interest. JPEG2000 Part 2 allows the employment of custom WPBs [15], [16], but WPBs for JPEG2000 have not been subject to extensive investigations so far. While wavelet-packet selection has been investigated for several proprietary compression systems, a thorough discussion for JPEG2000 is still missing, which is the only standardized and widely available compression system that offers the application of custom wavelet packets in its Part 2.

In [17], the variants of representing WPBs as discussed during the development of the JPEG2000 Part 2 standard have been assessed with respect to compression performance. For image confidentiality, it has been proposed to use secret wavelet packet bases as a means for compression integrated JPEG2000 encryption [18] (where the impact on compression performance needs to be controlled). Interestingly, at least to the best of the authors' knowledge, optimal wavelet packet basis selection in a rate-distortion sense [4] has not been discussed for JPEG2000 so far.

In this work we show that efficient, best WPB selection is possible in JPEG2000 by an extension of the approach of [4]. We define (and develop an algorithm for) the Lagrangian cost of a subband in JPEG2000, which enables the determination of the best WPB in a rate-distortion sense. Thus for the first time, the maximum performance gains achievable by an optimal selection of WPBs in JPEG2000 can be assessed. The influence of header data on rate-distortion optimal WPB (RDO-WPB) selection is analyzed and evaluated in-depth. Our focus in this work is on highly textured image data, especially on fingerprint images, for which a custom WPB has been proposed. Additionally, the computational complexity of the best WPB selection algorithms for JPEG2000 is discussed and compared to the classical dyadic decomposition, as mandatory in JPEG2000 Part 1.

There are substantial extensions to own previous work [19], [20] as well: The development, implementation and evaluation of a concise header cost determination algorithm, the implementation and evaluation of a lossless coding mode for our rate-distortion optimal wavelet packet coder, and the improvement of the evaluation framework, which now analyzes the compression performance with state-of-the-art quality metrics. A novel packet header cost determination algorithm has been developed, which allows to assess the actual cost of the packet header portion of a subband, thus enabling rate-distortion optimization without the imprecisions of the header cost estimation. Thus our JPEG2000-based WPB coder achieves always better or equal results compared to the underlying JPEG2000 Part 1 coder. This is further the first contribution that specifically

discusses rate-distortion optimal anisotropic wavelet packet selection for JPEG2000. The source code of our coder and evaluations is publicly available.¹

Section II gives an overview of JPEG2000, and Section III discusses algorithms for rate-distortion optimal wavelet packet basis selection within JPEG2000. Anisotropic wavelet packet selection is presented in Section III-B and the complexity of both isotropic and anisotropic rate-distortion optimal WPB selection is analyzed in Section IV. More efficient wavelet packet basis selection with computationally efficient heuristics is discussed in Section V. Section VI presents experimental results on fingerprint databases and other textured data. We draw our final conclusion in Section VII.

II. OVERVIEW OF JPEG2000

JPEG2000 employs a wavelet transform and uses the EBCOT-algorithm (embedded block coding with optimized truncation) to encode the wavelet coefficients. The wavelet coefficients of a subband are grouped in rectangular blocks (codeblocks), which are coded independently to separate bitstreams. JPEG2000 Part 2 [21] allows arbitrary WPBs. The standard [21, p.54] restricts the set of permissible WPBs, every high-frequency subband may only be decomposed two more times (vertically, horizontally or both). In Fig. 1, the WSQ-WPB is in the set of permissible WPBs, while the best bases for the Artificial and the Barbara image are not. In this work we evaluate whether these restrictions of Part 2 have a negative impact on compression performance.

A JPEG2000 file (codestream) consists of a main header followed by several packets. Each packet increases the decoded image quality. Each packet belongs to a certain quality layer and resolution. The number of quality layers can be freely chosen (for the scope of this work we set the number of quality layers to one). A packet consists of a packet header and a packet body. The packet body is solely comprised of bitstreams (coded codeblock data). The packet header contains information necessary for interpreting and decoding packet body data. The following data is written in the packet header for each codeblock of the subbands of the packet's resolution: leading zero bitplanes, the length of codeblock contribution, the number of coding passes and the inclusion information. The packet header formation is illustrated in Fig. 3. II denotes inclusion information, i.e., whether the code block contributes to the packet. LZB denotes leading zero bitplanes of the coefficients of a codeblock. NCP denotes the number of contributing coding passes (EBCOT employs three coding passes for a single bitplane of the coefficients of a codeblock), and CCPL denotes the length of the coded codeblock contribution in the packet body.

A. Rate-Distortion Optimization in JPEG2000

The embedded bitstream of a single codeblock has several potential truncation points, i.e., each codeblock has a separate RD function. The goal of an encoder is to arrange the bitstream data of all codeblocks in an RD optimal manner, i.e., to find the truncation points that minimize the distortion for a given rate. The most common algorithm for JPEG2000 is post-compression-rate-distortion optimization (PCRD-Opt). Fig. 4 illustrates

¹<http://www.wavelab.at/sources/>

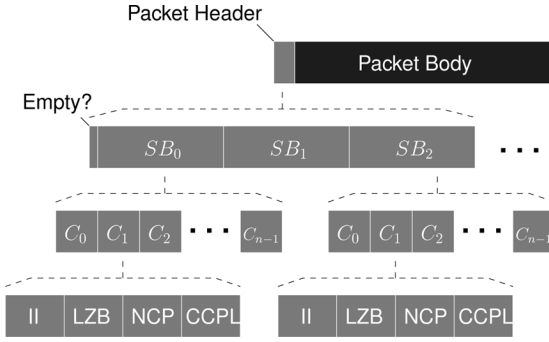


Fig. 3. Packet header formation.

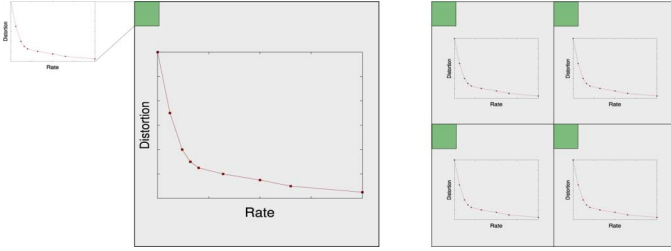


Fig. 4. Rate-distortion statistics of codeblocks and a subband and its children: the isotropic case.

the basic setup: A subband consists of one to several codeblocks (illustrated by green squares), each with its own rate-distortion statistic. An optimally coded JPEG2000 codestream can be obtained by selecting an individual rate for each codeblock. A truncation point of the codeblock B_i is denoted by n_i , all truncation points by \vec{n} . The embedded bitstream of the codeblock B_i can be truncated to a rate $R_i^{n_i}$ (for a given truncation point n_i). The rate constraint is then

$$R(\vec{n}) = \sum_{i=1}^{\#(\text{cbs of image})} R_i^{n_i} \leq R_{\max}. \quad (1)$$

The distortion of each codeblock B_i for a truncation point n_i is given by $D_i^{n_i}$. Given an additive distortion measure, the distortion D of the compressed image is derived by

$$D(\vec{n}) = \sum_{i=1}^{\#(\text{cbs of image})} D_i^{n_i}. \quad (2)$$

An optimal solution (minimizing D) of truncation points \vec{n}^* for this constrained problem can be found by solving the corresponding unconstrained problem (Lagrangian RDO):

$$\vec{n}^* = \arg\min_{\vec{n}} [D(\vec{n}) + \lambda R(\vec{n})]. \quad (3)$$

Considering D a function of R , a solution is obtained by setting $D'(R) = -\lambda$, which yields $D'_i(R) = -\lambda$.

Note that JPEG2000 RDO by PCRD-Opt requires that the distortion measure is additive across subbands, e.g., the distortion of a subband can be computed by adding the distortion of its children. This requirement is met only approximately by the wavelet transforms of JPEG2000 Part 1.

The rates (length) of the codeblock contributions need to be coded as well, this information is contained in the packet

headers. In the PCRD-Opt algorithm the packet header data is taken into account after the determination of the rates of the codeblock contributions, a method shown to perform robustly in JPEG2000 codestream assembly [15]. Note that packet headers are always coded for an entire resolution, i.e., for all codeblocks for all subbands (LL or HL, LH, and HH) of a resolution.

III. BEST WAVELET PACKET BASES

For compression, the best WPB is the one that minimizes the size of the compressed image at a given level of distortion (best WPB in a rate-distortion sense). Thus finding the best solution in a rate-distortion sense depends on the underlying coding mechanisms and might be computationally complex and complex to integrate in a compression framework such as JPEG2000.

Testing every possible WPB soon becomes infeasible, as the number of possible WPBs grows tremendously with the decomposition depth d . The following recursion [22] calculates Q_d , the number of possible isotropic WPBs at depth j :

$$Q_d = Q_{d-1}^4 + 1 \quad (4)$$

where $Q_0 = 1$. At depth two we have 17 possible WPBs, at depth three 83 522, at depth four 4.9×10^{19} , at depth five 5.6×10^{78} , at depth six 9.9×10^{314} , and at depth seven 9.6×10^{1259} .

There is a more efficient algorithm for the determination of the “best” WPB (best in the restricted sense of the employed cost function only): the best basis algorithm (BBA) [1]. The BBA first generates the full wavelet packet decomposition tree at maximum decomposition depth and starts from the leaves, i.e., the subbands at the deepest decomposition depth. The root node of the wavelet packet decomposition tree corresponds to the entire image and each tree node (except the leaves at maximum decomposition depth) has four children (the LL, HL, LH, HH subbands). Fig. 4 illustrates a root subband (with its own rate-distortion statistic) and its four children (with their own rate-distortion statistics). An encoding of a wavelet packet decomposition tree visits the tree in a depth first scan, emitting a “1” if subband is further decomposed and a “0” otherwise. The BBA merges the children subbands of a parent subband if the sum of the costs of its children is higher than the parent’s cost. A prerequisite for the BBA is that the cost function is additive, i.e., let \mathcal{S} denote a wavelet decomposition tree and n_i denote its terminal nodes (subbands that are actually coded). Then a cost function c is called additive if

$$c(\mathcal{S}) = \sum_i c(n_i).$$

A. Best WPBs in a Rate-Distortion Sense

The best solution for an actual compression framework in an RD sense is obtained if the coding costs are not estimated, but actually determined. Prerequisites for the RD-optimal wavelet packet selection algorithm are that the nodes are coded independently and that the distortion measure is additive. If these prerequisites hold, then the Lagrangian cost function is additive and the BBA algorithm is guaranteed to find the best basis in terms of compression performance, i.e., the WPB with the

lowest distortion for a certain target rate. The cost of a subband sb is calculated by the Lagrangian cost function which is defined as

$$J(\lambda)_{sb} = D_{sb,\lambda} + \lambda R_{sb,\lambda}. \quad (5)$$

Children subbands are not merged if the following split condition holds:

$$J(\lambda)_{parent} \geq \sum_{child=1}^{\#(children)} J(\lambda)_{child}. \quad (6)$$

In order to obtain a solution for a target bitrate, an efficient bi-section search on the parameter λ can be conducted [4].

1) *Lagrangian Cost of a Subband for JPEG2000*: The essential part of integrating the algorithm for finding the best WPB in a rate-distortion sense into JPEG2000 is to appropriately determine the Lagrangian cost of a subband.

RDO WPB selection requires that the subbands are coded independently: this requirement is met by the coding framework of JPEG2000 as we will subsequently show in greater detail. Additionally, RDO WPB selection requires that the distortion measure is additive, i.e., the distortion of a wavelet decomposition tree can be computed by adding the distortion of its terminal nodes (subbands). The distortion of a wavelet transform is additive if the transform is orthonormal or the quantization errors are uncorrelated [15]. Although JPEG2000 Part 2 allows the application of custom wavelet transforms, including orthonormal wavelets, the 9–7 and 5–3 wavelet transforms of Part 1 are “only” biorthogonal, which does not strictly guarantee additive distortion computation. However, distortion additivity is also prerequisite for JPEG2000 RDO [15] with the PCRD-Opt algorithm, i.e., the entire JPEG2000 Part 1 distortion optimization works on the basis of the approximate distortion additivity. The approximate distortion additivity is argued to be sufficient for practical RDO [15]. In the following, we will treat the distortion as additive, an approach further justified as orthonormal wavelets, e.g., the Haar wavelet, can be employed in Part 2 and in our custom implementation as well. However, the Haar wavelet performs substantially worse than the irreversible 9–7 transform in terms of compression performance, which is also the case for wavelet packet bases.

A subband consists of several codeblocks, each with a bitstream with its own rate-distortion statistics, i.e., truncation points and associated distortions. These data has to be used to compute the optimal rate-distortion function for a subband, i.e., the operation points with the least distortion for a rate. The rate-distortion function of subband is derived by Lagrangian optimization as well.

Given the rate-distortion statistics of the codeblocks of a subband, an optimal solution is found for a certain value of the parameter λ by selecting all the truncation points for which the absolute slope of the rate-distortion function is larger than or equal to the rate-distortion threshold λ (see Algorithm 1). As the codeblocks only have a discrete number of truncation points and associated slopes it is sufficient to consider these slopes in the rate-distortion optimization. This Lagrangian optimization calculates the rate-distortion function of a subband; however, this optimization ignores the signalling overhead necessary to code the length information of the truncation points.

This length information is contained in the JPEG2000 packet headers, which are coded for each resolution, i.e., commonly multiple subbands. Thus this optimization minimizes the overall packet body size, and only minimizes the overall file size if the cost of coding the headers is not influenced by the selection of the WPB (we refer to this algorithm as RDO-WPB).

The Lagrangian cost of a subband can be computed as the sum of the Lagrangian costs of its codeblocks:

$$\begin{aligned} J(\lambda)_{sb} &= \sum_{cb=1}^{\#(cbs \text{ of subband})} J(\lambda)_{cb} \\ &= \sum_{cb=1}^{\#(cbs \text{ of subband})} D_{cb,\lambda} + \lambda R_{cb,\lambda}. \end{aligned} \quad (7)$$

The actual algorithm to determine the Lagrangian cost of a subband is given in pseudo-code (see Algorithm 1).

Algorithm 1: Lagrangian cost function of a subband

Param: λ

costs = 0

for ($b = 0$; $b < \#(\text{code-blocks of the subband})$; $b++$) **do**

for ($\text{slopeIdx} = 0$; $\text{slopeIdx} < \#(\text{absSlopes of codeblock}[b])$; $\text{slopeIdx}++$) **do**

 absSlope = blocks[b].absSlopes[slopeIdx]

if absSlope $< \lambda$ **then**

 break;

end if

end for

 costs+ = getDistortion(slopeIdx - 1) + lambda * getRate(slopeIdx - 1)

end for

return costs

2) *Considering the Packet Header in the Lagrangian Cost of a Subband*: In Fig. 5(a) the packet header cost for the LL subband is analyzed in detail, the cost of the packet headers is plotted for increasing decomposition depths and varying rate. The LL subband’s packet header cost is compared to the other subbands in Fig. 5(b). The packet header cost increases monotonically (almost linear) with the overall target bitrate in these examples. In the RDO-WPB algorithm the (packet) header data cost is considered constant and independent of the decomposition, i.e., the packet header cost of a subband and a further decomposed subband (the sum of the packet header costs of its children) are assumed to be equal: A simplification with unpredictable and potentially disastrous consequences for the performance of RDO wavelet packet selection.

Although packets and thus packet headers are written per resolution, i.e., for several subbands, the packet header data is coded per subband, each subband maintains its own coding

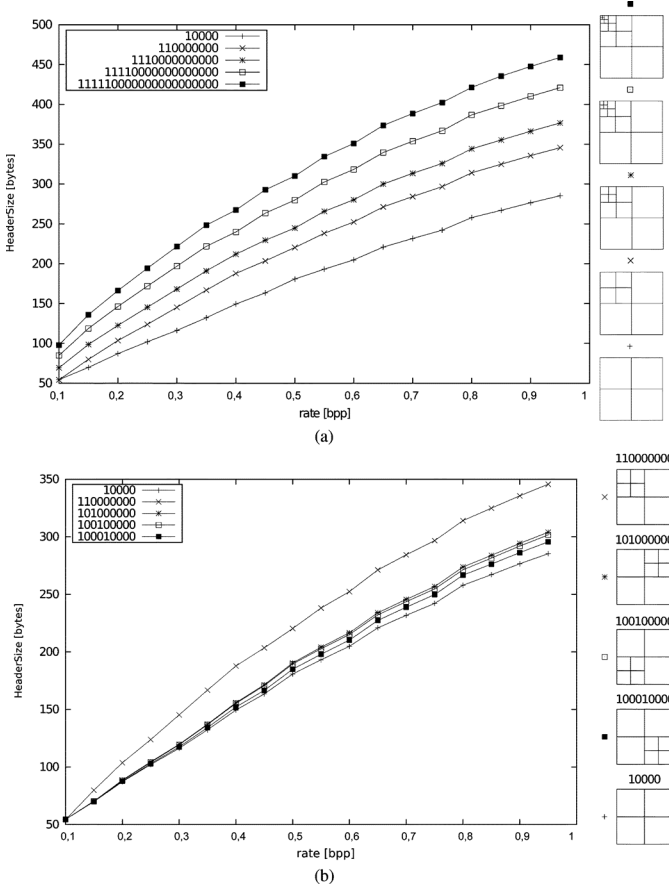


Fig. 5. Packet header data for different decomposition depths and subbands. (a) Splitting the LL subband. (b) Comparison of the costs of different subbands.

states, e.g., in the form of associated tag trees. This independent per subband coding of the packet header portion is a prerequisite for RDO wavelet packet selection [4]. An optimal rate-distortion function for the packet body data of a subband, i.e., the arithmetically coded coefficient data of the codeblocks, can be obtained by Lagrangian optimization on the basis of its codeblocks rate-distortion statistics. The optimal rate-distortion performance points for a subband can be efficiently obtained by the Lagrangian rate-distortion optimization of the subband's packet body data, by simply adding the packet header rate to the rate of the packet body data, which will result in an optimal result if the packet header size behaves "reasonably" as function of the packet body size. Although packet header coding is quite a complex context adaptive procedure, our empirical results clearly indicate an approximately linear growth [see Fig. 5(a) and (b)].

Thus the most concise RDO WPB selection algorithm builds individual rate-distortion statistics for each subband in which the subband's packet header data is included. The subband's RD slopes are computed after the Lagrangian rate-distortion optimization of its packet body data and the simulated coding of the corresponding packet header portion. These RD statistics of the subbands are then employed for the RDO basis selection algorithm. This entire wavelet packet basis selection algorithm is referred to as RDOH WPB. If the assumptions distortion additivity and monotonicity of the packet header data are met this algorithm provably selects the best wavelet packet basis.

For each subband and lambda the optimal truncation points for the coded code block data are determined and afterwards the packet header portion of the subband is determined by simulating the coding in dependency of the packet body RDO, i.e., the obtained truncation points. Thus we propose to employ the following Lagrangian cost of a subband sb in dependency of the Lagrangian parameter λ which determines the rate-distortion points of a subband in dependency of the newly computed slope (with inclusion of the packet header data):

$$J(\lambda)_{sb} = D_{sb,\lambda} + \lambda(R_{sb,\lambda} + R_{sb,\lambda}^H). \quad (8)$$

The header cost has to be integrated in the Lagrangian cost of a subband in order to reflect the increased coding cost for a further decomposition of a subband. In [19] it is proposed to estimate the actual header cost of a subband. The Lagrangian cost of a subband sb at depth d is computed as follows in [19]:

$$J(\lambda)_{sb} = D_{sb,\lambda} + \lambda(R_{sb,\lambda} + R_{sb}^H). \quad (9)$$

We refer to this extended algorithm by RDOE WPB and employ a very simple packet header cost estimation, i.e., a fixed cost for coding a single codeblock and a single subband. Among others, the packet header cost for subbands and further decomposed subbands depends on the triple: size of the image (x, y) , code-block size (cb_x, cb_y) , and wavelet decomposition depth of the subband under investigation d . If $cb_x > x/2^d$ or $cb_y > y/2^d$, the subband is decomposed into subbands smaller than the codeblock size and additional entries for the new codeblocks have to be added to the packet header, and the packet header length is increased. E.g., for 512×512 images, and $d = 5$, further decomposing the subbands at depth 3 becomes more expensive in terms of the number code blocks. We assign a fixed cost for codeblocks that result from too deep decompositions that result in smaller subbands than codeblocks. Additionally we consider the quantization type which impacts the header as well, i.e., the signalling cost of quantization type specific is added to the subband cost. For quantization type expounded the signalling cost is 16 (2 bytes in the main header) and for quantization type reversible the extra cost is 8 (1 byte in the main header) [23].

The simplification of RDO-WPB has to be paid by sub-optimal compression performance, as can be seen in Fig. 6, where RDO-WPB is clearly outperformed by the algorithm RDOH WPB, which takes header data into account. Considering only the packet body size, RDO-WPB is optimal (see Fig. 7). The performance gains for RDOH WPB for the "Artificial" image are enormous, and considerable for the "Barbara" image as well (see Figs. 8 and 9). For both RDOH and RDOE outperform RDO-WPB.

3) *Bitrate Adjustment: A Bisection Search for the Lagrangian*
 λ : The actual bitrate of the best basis algorithms in a JPEG2000 rate-distortion sense is determined by the parameter λ . Thus in order to achieve a certain target bitrate the appropriate value of λ has to be determined. Our implementation employs a classic bisection method [4], [24]. Starting from a lower (λ_{\min}) and an upper bound (λ_{\max}) for λ , i.e., (10) has to be satisfied, a bisection search is performed, as outlined in algorithm 2:

$$R(\lambda_{\min}) \leq R_{budget} \leq R(\lambda_{\max}). \quad (10)$$

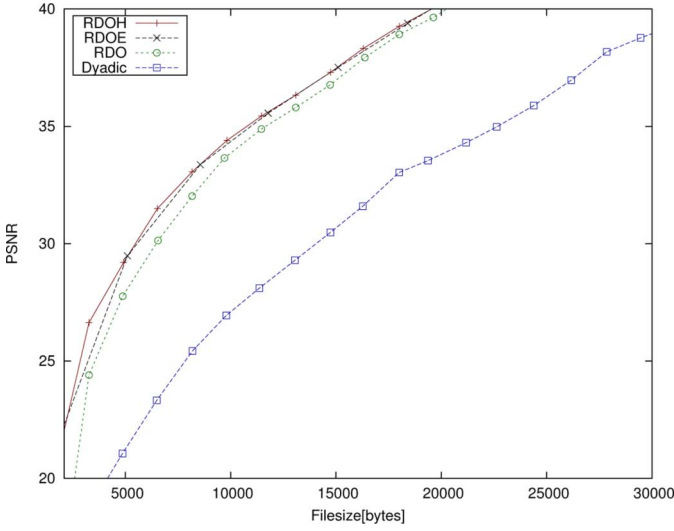


Fig. 6. Artificial image: file size.

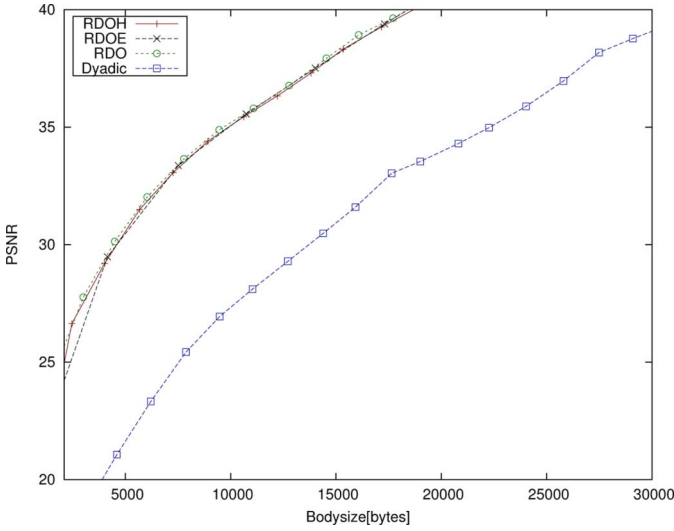


Fig. 7. Artificial image: body size.

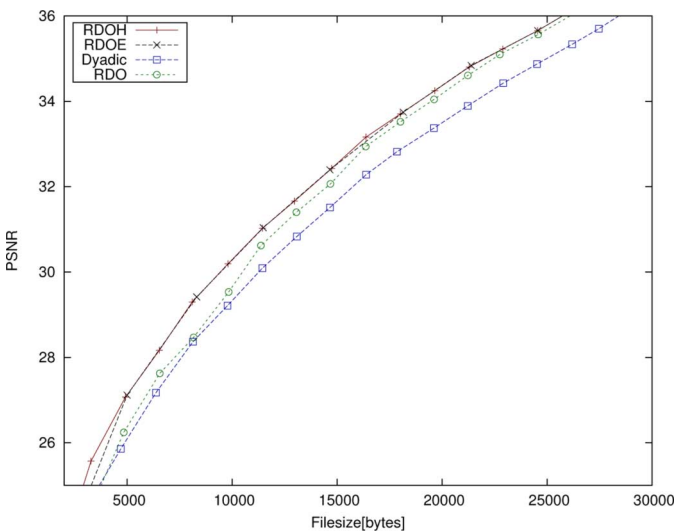


Fig. 8. Barbara image: file size.

Algorithm 2: Bisection search for λ

Param: R_{budget}

$$\lambda = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})/2$$

$$R_{result} = 0$$

$$\text{maxSearchSteps} = 32$$

$$\text{numSearchSteps} = 0$$

$$\epsilon = 10^{-6}$$

While $\text{numSearchSteps} < \text{maxSearchSteps}$ **do** $\text{numSearchSteps}++$ $R_{current} = \text{getRateForLambda}(\lambda)$ **If** $R_{current} == R_{budget}$ **then**

$\lambda_{result} = \lambda$

$R_{result} = R_{budget}$

break; **end if** **if** $R_{current} < R_{budget}$ **then** **if** $R_{result} < R_{current}$ **then**

$\lambda_{result} = \lambda$

$R_{result} = R_{current}$

end if **if** $|\lambda_{\max} - \lambda| < \epsilon$ **then** **break;** **end if**

$\lambda_{\max} = \lambda$

$\lambda = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})/2$

else if $R_{current} > R_{budget}$ **then** **if** $|\lambda_{\min} - \lambda| < \epsilon$ **then** **break;** **end if**

$\lambda_{\min} = \lambda$

$\lambda = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min})/2$

end if**end while****return** λ_{result}

The initial values of λ_{\min} and λ_{\max} can be set to the minimum and maximum slope of the rate-distortion functions of all codeblocks; suitable initial values have been found experimentally to be 0 and 10.

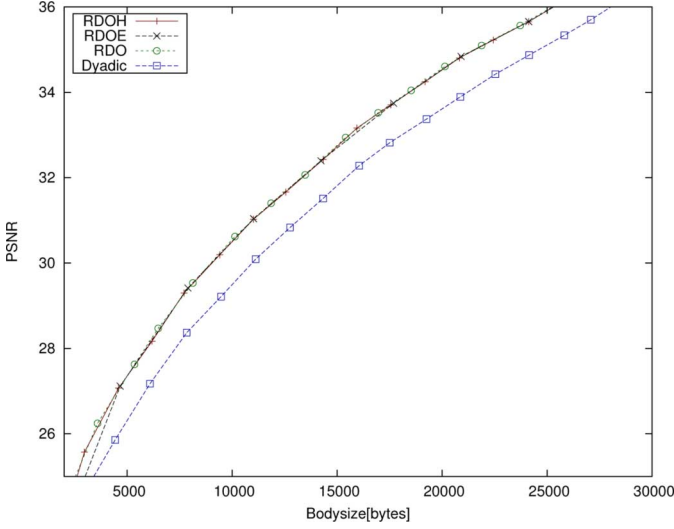


Fig. 9. Barbara image: body size.

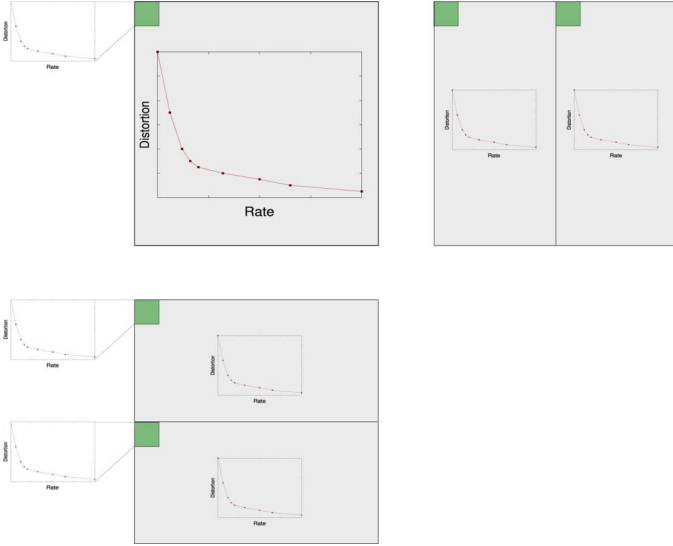


Fig. 10. Rate-distortion statistics of codeblocks a subband and its potential children: the anisotropic case.

B. Best Anisotropic Wavelet Packet Bases

Anisotropic wavelet packet bases allow to decompose a subband horizontally or vertically, while isotropic wavelet packet bases only allow a decomposition in both direction, i.e., horizontally and vertically simultaneously. Fig. 10 illustrates a subband and its two possible decompositions (vertical and horizontal). Isotropic wavelet packets can be modelled as an ensemble of unique wavelet packet trees, i.e., there is exactly one decomposition tree for each decomposition structure/WPB, which is not the case for anisotropic WPBs. The root of an anisotropic wavelet packet decomposition tree represents the image, i.e., the root subband. A subband has either no or two horizontal children or two vertical children, which are derived by horizontal or vertical wavelet decomposition. For anisotropic WPB the ensemble of (decomposition) trees is not unique, i.e., several decomposition trees describe the same wavelet packet basis. E.g., vertical decomposition and horizontal decomposition of the two children is equivalent to

horizontal decomposition and vertical decomposition of the two children, which corresponds to a single decomposition in the isotropic case. Note that the anisotropic decomposition tree of an isotropic WPB has twice the depth of the corresponding isotropic decomposition tree, i.e., the anisotropic depth is twice the isotropic depth. An encoding of an anisotropic wavelet packet tree visits the tree in a depth first scan (visiting first nodes corresponding to the low pass subbands) and emitting a “1” if a subband is decomposed (a “0” otherwise), followed by a “1” if the subband is decomposed horizontally and a “0” if the subband is decomposed vertically. There are substantially more anisotropic decomposition trees than anisotropic wavelet packet bases. The number of anisotropic wavelet packet bases is given by the following recursion ($A_{-2} = A_{-1} = 0$):

$$A_d = 1 + 2A_{d-1}^2 - A_{d-2}^4.$$

At an anisotropic depth of 1 there are 3 possible anisotropic WPB (no decomposition, vertical horizontal), at an anisotropic depth of 2 there are already 18 bases (for the comparable isotropic depth of 1 we have only 2 possible isotropic WPB), at an anisotropic depth of 4 there are 540 273 anisotropic WPB (compared to 17 isotropic WPB for the comparable isotropic depth of 2), at an anisotropic depth of 6 there are $\approx 3.75 \times 10^{23}$ bases (isotropic: 83 522), and at an anisotropic depth of 8 there are $\approx 8.4 \times 10^{94}$ (isotropic: $\approx 4.9 \times 10^{19}$). Thus there are substantially more anisotropic WPB than isotropic WPB for a comparable decomposition depth.

For the RDO anisotropic WPB selection we need to clarify the parent child relationship. A subband can be decomposed either horizontally or vertically, thus has two types of children. Thus in contrast to isotropic WPB selection, where only one decision had to be made (merge or split the subband), in the anisotropic case there are three options: merge, split horizontally or split vertically. Apart from this distinction the basic tree pruning algorithm works similar as for the isotropic case.

The major difference is in terms of complexity (which is discussed in detail in Section IV), the anisotropic RDO-WPB selection is substantially more complex compared to isotropic WPB selection. This is reflected in its computational complexity estimate which is in $\mathcal{O}(d^2 \times n)$ (even the most efficient implementation), while isotropic WPB selection is in $\mathcal{O}(d \times n)$. Fig. 11 illustrates the necessary subbands for best basis selection for a isotropic decomposition depth of 3 which corresponds to an anisotropic decomposition depth of 6.

Anisotropic wavelet packet basis can be restricted by the decomposition depth which we perform jointly, i.e., the number of horizontal plus the number of vertical decompositions must not be in excess of the maximum decomposition depth.

IV. COMPLEXITY OF RATE-DISTORTION OPTIMAL WPB SELECTION

The asymptotic complexity of isotropic RDO WPB selection for a maximum decomposition depth d and an n -element signal is in $\mathcal{O}(d \times n)$ and also in $\mathcal{O}(n \log n)$, because d is bounded by $\log n$, which is the maximal decomposition depth. The asymptotic complexity of anisotropic RDO WPB selection for a maximum decomposition depth d and an n -element signal is in $\mathcal{O}(d^2 \times n)$ and also in $\mathcal{O}(n \log^2 n)$. Specifically,

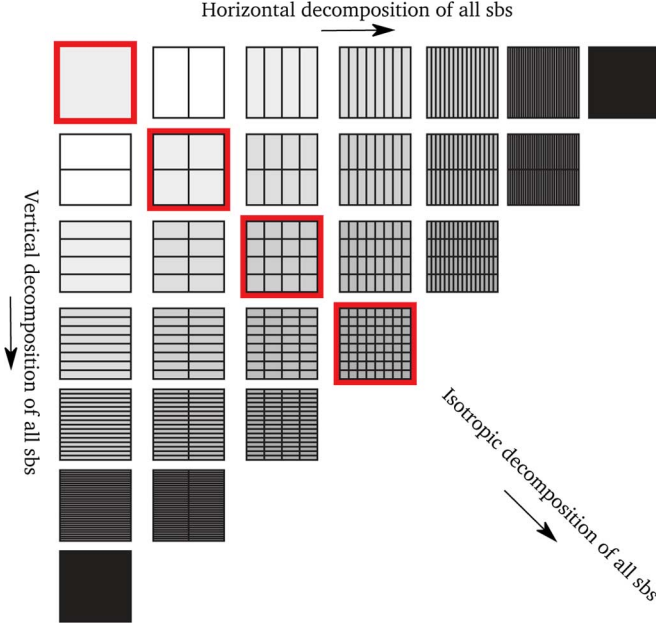


Fig. 11. Necessary subbands of isotropic (red) versus anisotropic best basis selection.

for isotropic RDO WPB selection $d + 1$ JPEG2000 images have to be generated, while for anisotropic RDO WPB selection $(d + 1) \times d/2$ JPEG2000 images have to be generated (referred to as full JPEG2000 compressions). Fig. 11 illustrates the WPBs that have to be fully JPEG2000 compressed in order to obtain all subbands, e.g., the complexity of RDO WPB selection with $d = 5$ is roughly the same as the complexity of anisotropic RDO WPB selection with $d = 2$. The complexity of anisotropic RDO WPB selection with $d = 3$ already costs 10 full JPEG2000 compressions, and 1.66 times the complexity of isotropic RDO WPB selection with $d = 5$.

In order to assess the concrete complexity of isotropic RDO WPB we consider the computationally complex parts of the JPEG2000 compression pipeline: DWT (and quantization), coding of codeblocks (and RDO), as well as file I/O (including final bitstream formation). Depending on the implementation, the compression settings and source data the shares vary; for JPEG2000 Part 1, overall 0.8 s are needed for a 512×512 image with JJ2000 default settings (all evaluations are performed on an Intel Core2 6700@2.66GHz). For the isotropic RDO WPB with a maximum decomposition depth d a full wavelet decomposition for every depth l ($1 \leq l \leq d$) has to be performed. The coefficients of a full decomposition at depth l can be used to compute the coefficients of the next depth $l + 1$. In terms of a DWT at depth 1, \mathcal{W} , the cost of all decompositions is at least $d \times \mathcal{W}$. However, in practice the runtime complexity is tremendously increased for decomposition depths greater 5 (see Table II). Coding and RDO has to be done at every depth l and for no decomposition as well, which adds up to $(d + 1) \times \mathcal{B}$, where \mathcal{B} is the cost of coding all coefficients. As long as the subbands are larger than the codeblocks, their coding and RDO cost remain approximately constant for all depths.

If the subbands become smaller than the codeblocks, the runtime performance decreases; however, this effect is implementation-specific (see Table I for JJ2000's behavior).

TABLE I
RUNTIME PERFORMANCE IN SECONDS DEPENDING ON CODEBLOCK SIZE FOR 512×512 IMAGES AND NO DWT

$cb_x cb_y$	64^2	32^2	16^2	8^2	4^2
No DWT	0.6s	0.7s	0.9s	1.5s	3.4s

TABLE II
RUNTIME PERFORMANCE IN SECONDS DEPENDING ON DECOMPOSITION DEPTH d OF FULL DECOMPOSITION (FD) AND RDO[H]-WPB (WPB) FOR 512×512 IMAGES

d	0	1	2	3	4	5	6
$cb_x cb_y$	64^2	64^2	64^2	64^2	32^2	16^2	8^2
FD	0.6s	0.63s	0.7s	0.75s	0.9s	1.8s	15s
WPB	0.6s	1.00s	1.1s	1.30s	1.5s	2.8s	18s

The overall cost for isotropic RDO WPB at depth d , $\mathcal{R}(d)$, in terms of a compression at depth 1, \mathcal{C} , is approximately $\mathcal{R}(d) \approx d \times \mathcal{C} - (d - 1) \times \mathcal{D}$, where \mathcal{D} represents fixed time, e.g., for the actual JJ2000 implementation: Java start up time, image IO, and bitstream IO (approx. 0.3s).

In conclusion, for a reasonable wavelet decomposition depth of 4 our isotropic RDO WPB implementation only takes less than twice the default JJ2000 compression time, which is in-line with our theoretical analysis, which predicts $4 \times 0.63s - 3 \times 0.3s = 1.62s \approx 1.5s$.

The complexity of RDOH is increased by the coding of the packet header portions of a subband, which becomes significant for higher decomposition depths as the number of subbands n_{sb} grows exponentially with the decomposition depth d of a full wavelet packet decomposition ($n_{sb} = 4^d$). At a decomposition depth of 5 a single RDOH compression takes about double the time of the RDOE version with the packet header estimate. Thus for higher decomposition depths the estimation of the packet header cost is the method of choice.

V. MORE EFFICIENT WAVELET PACKET BASIS SELECTION

For practical application efficient wavelet packet basis selection is fundamental. An approach towards cutting the cost of basis selection is to refrain from using the actual coding costs (as done in RDO WPB selection) and employ computationally more efficient cost functions. If these cost functions are additive then the determined basis is optimal in the sense of the cost function. However, this optimality in terms of a cost function does not necessarily go hand in hand with optimality in terms of RD performance, our main goal in the scope of this work. The BBA with cost functions still is computationally complex, especially for anisotropic WPBs, as all the subbands have to be computed, i.e., the full decomposition tree has to be computed, which is then visited in a bottom up fashion. The computational complexity of the anisotropic BBA is in $\mathcal{O}(d^2 \times n)$, while the isotropic BBA is in $\mathcal{O}(d \times n)$. For the BBA a cost for every possible subband has to be computed, the number of subbands is far more in the anisotropic case than in the isotropic case (see Fig. 11).

Computational complexity is significantly reduced if the decomposition tree is visited in a top-down fashion and if the subbands are only computed if necessary. This algorithm is briefly sketched:

- Decompose a subband (in case of anisotropic WPBs horizontally and vertically).
- Compute costs of the subband and its children.
- If subbands cost is minimal, stop; else evaluate each child of the minimum branch.

This top-down algorithm does not guarantee finding an optimal basis for an additive cost function, however, the optimality in terms of a cost function does not imply optimality in a rate-distortion sense anyway.

A. Cost Functions

Alternatively to the optimal wavelet packet basis in a rate-distortion sense with the actual coding bitrate as cost function it has often been proposed to employ simpler cost functions for best basis selection (although these may not result in best bases in a rate-distortion sense). In this section we will present common cost functions. Let c_i represent the value of the coefficients of a subband. The following additive cost functions are calculated:

- L1-norm: $\sum_i |c_i|$
- L2-norm: $\sum_i c_i^2$
- LogE-log energy metric: $\sum_i \ln(c_i^2)$.
- EIC—entropy information cost or Shannon metric [1]: $\sum_i c_i^2 \ln(c_i^2)$

Furthermore, we employ an entropy based cost function, which basically computes an entropy estimate for the quantized coefficients of a subband. The coefficients are quantized, i.e., divided by q and rounded to the next integer, their distribution statistics are calculated and these data is used to compute an entropy estimate. This entropy estimate is weighted with/multiplied by the number of coefficients in the subband. We refer to this cost function as weighed entropy estimate, E_q , where q indicates the divisor in the quantization process. We have extended the basic entropy cost by a penalty for subbands that become smaller than the codeblock size (a constant $k = 2$ is added for every codeblock contained in the subband, which can be interpreted as two bit extra coding cost). Additionally we consider the JPEG2000 specific cost of signalling the quantization information for a subband, i.e., the signalling cost is added to the subband cost. For quantization type expounded the signalling cost is 16 (2 bytes in the main header) and for quantization type reversible the extra cost is 8 (1 byte in the main header) [23]. These simple and computationally inexpensive extensions model JPEG2000 coding specifics quite well. Overall, our entropy based cost function is computed very efficiently using hash tables.

B. Complexity

Complexity is reduced to $\mathcal{O}(\bar{d} \times n)$ for both anisotropic and isotropic WPB selection, where \bar{d} denotes the average decomposition depth of a subband, which corresponds to the average number of filter operations necessary to obtain a coefficient (two filter operations in the isotropic case and one filter operation in the anisotropic case). It can be obtained by the following algorithm (see Algorithm 3), that takes the encoding of an isotropic

wavelet packet tree and computes its average decomposition depth. The algorithm for anisotropic wavelet packet trees is similar (see Algorithm 4).

Algorithm 3: Average decomposition depth of an isotropic wavelet packet tree

function D(s)

b = s.pop_front() ▷
Get the leading bit and remove it from the bitstring.

if b == '0' **then**

return 0

else if b == '1' **then**

return 1 + (D(s) + D(s) + D(s) + D(s))/4

end if

end function

Algorithm 4: Average decomposition depth of an anisotropic wavelet packet tree

function A(s)

b = s.pop_front()

if b == '0' **then**

return 0

else if b == '1' **then**

s.pop_front()

return 1 + (A(s) + A(s))/2

end if

end function

The main impact on the complexity of all top-down approaches is the average decomposition depth, \bar{d} , of the selected WPB, which is highly source image dependent (see Fig. 1), e.g., $\bar{d} \approx 1.33$ for the Lena image. JPEG2000 Part 2 imposes an upper bound for \bar{d} of 3.3. The mean average decomposition depth for the bases selected by RDOH ($d = 5$) on the Brodatz database is remarkably low, on average $\bar{d} = 1.37$. An upper bound of the complexity of a top-down approach for an image with a WPB with \bar{d} is $C_T(\bar{d}) \leq (\bar{d} + 1) \times (\mathcal{W} + \mathcal{F}) + \mathcal{B} + \mathcal{D}$, where \mathcal{C} is the cost of JPEG2000 coding the image with one level wavelet decomposition and \mathcal{D} represents the fixed time, e.g., Java start-up time and file IO (timing measurements are given in Section IV). \mathcal{C} is mainly comprised of the wavelet transform cost, \mathcal{W} , the coefficient coding, \mathcal{B} , and \mathcal{D} . \mathcal{F} denotes the cost of evaluating the cost function on all coefficients. The performance improvements of a top-down approach rely on $\bar{d} + 1 \ll d$ and $\mathcal{F} \ll \mathcal{B}$.

In our implementation top-down best basis search complexity is almost the same as a single full JPEG2000 compression (efficient).

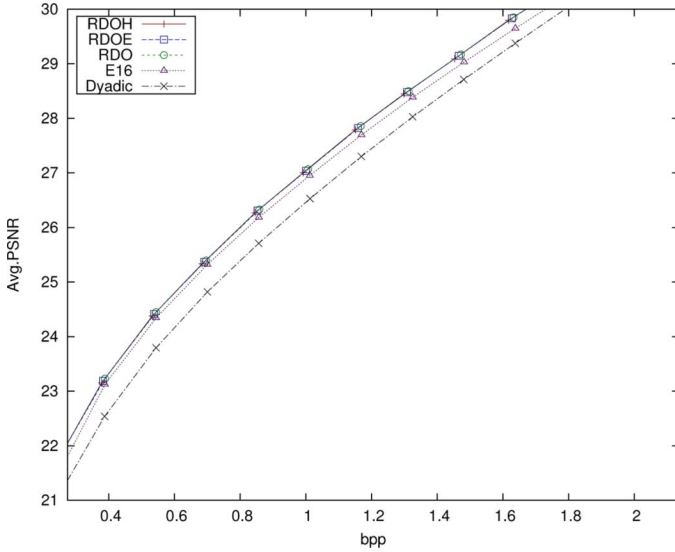


Fig. 12. Brodatz: PSNR evaluation.

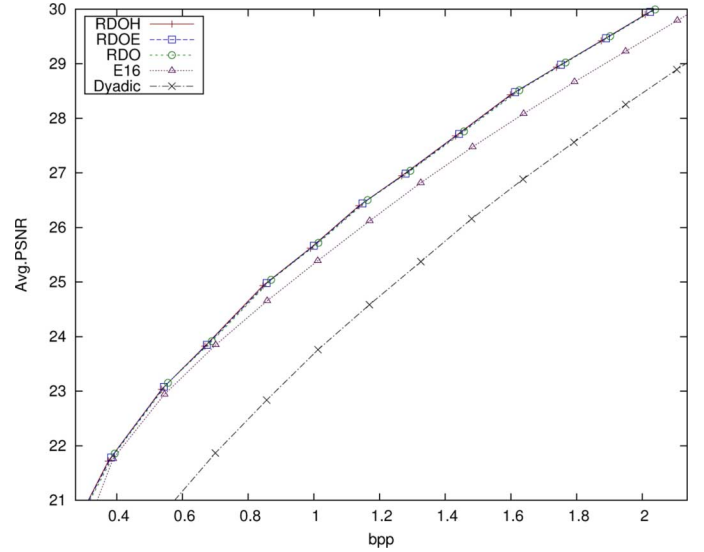


Fig. 13. A 20% subset of the Brodatz database: PSNR evaluation.

VI. EXPERIMENTAL RESULTS

The results have been produced with a custom implementation, which is based on the JJ2000 reference implementation. The correctness of our implementation of RDOH, RDOE and RDO-WPB has been experimentally verified for depth 2 and depth 3 by testing every possible WPB (for test images from the FVC04 database). For lossy compression the 9–7 irreversible filter with quantization type expounded has been employed and lossless compression the 5–3 reversible filter with quantization type reversible. A maximum decomposition depth of 5 (if not explicitly stated otherwise), one quality layer, and 64×64 code-blocks have been employed. Additionally to the well-known PSNR we have assessed the image quality with state-of-the-art metrics, such as the VIF, MSSIM, and SSIM [25]. The Matlab package `metrix_mux` has been employed.

We present results for highly textured data (Brodatz database) and for fingerprint data (FVC2004 database).

A. Comparison of RDOH, RDOE, and RDO-WPB

At higher decomposition depths RDOH is superior to RDO-WPB, i.e., the consideration of header data leads to performance improvements. For the Artificial image concise PSNR and VIF results are summarized in Figs. 6 and 7 for a maximum decomposition depth of 7. Enormous PSNR improvements are achieved for RDO, RDOE and RDOH (over 7dB), the RDOH and RDOE algorithms work reliable for all quality ranges and outperform RDO significantly, especially in lower quality range. The VIF results report a similar objective quality behavior. For the Barbara image improvements of about 1dB are achieved with RDOH and RDOE, which outperform RDO (see Fig. 8). The VIF shows a similar behavior, although smaller PSNR differences are no longer distinguishable. Thus at higher maximum decomposition depths the consideration of the header data is recommended, the estimation with RDOE works well and is negligible in terms of computational complexity.

On the Brodatz database and with a decomposition depth of 5, RDOH, RDOE, and RDO-WPB perform well (see Fig. 12),

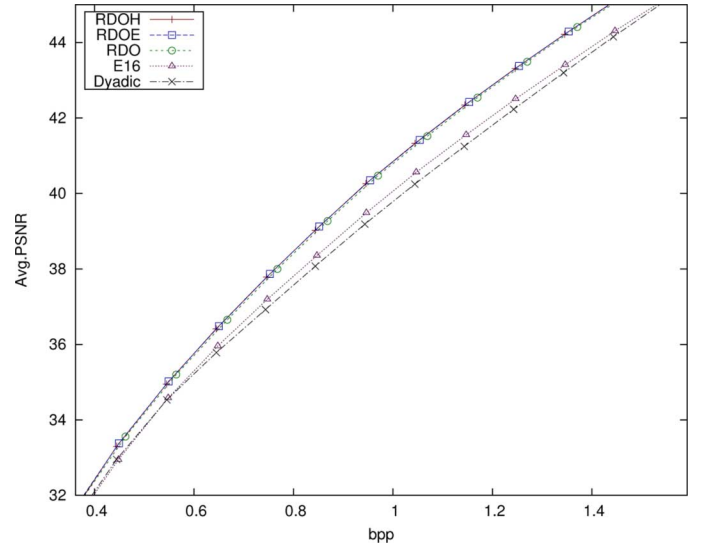


Fig. 14. FVC2004 DB4B: PSNR evaluation.

especially on a subset consisting of 20% of the Brodatz images (see Fig. 13). Thus for textured data, best basis selection in a JPEG2000 RDO sense can be recommended.

On fingerprint data RDOH, RDOE, and RDO-WPB selection also achieve a significantly improved performance; interestingly, the performance gains are more significant in the higher quality region starting a PSNR of 34 dB (see Fig. 14). The VIF evaluation supports the good performance for fingerprint data.

The best basis algorithm with our entropy based cost function (with quantization parameter $q = 16$) works very well on the Brodatz database (“E16” in Fig. 12). While the consideration of header data yields improvement for higher decomposition depths, the improvements are at a decomposition depth of 5 very small in average.

1) *Lossless Compression:* In Tables III–V the results for the test images, Artificial, Barbara, and Lena are given. In the first column the achieved size of the dyadic wavelet decomposition is given, the next column gives the reduction of the compressed

TABLE III
SELECTED TEST IMAGES WITH max. DECOMP. 5

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
artificial512	77953	20744	0	0
barbara	156784	235	0	1
lena	153529	18	0	1

TABLE IV
SELECTED TEST IMAGES WITH max. DECOMP. 6

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
artificial512	77931	20561	161	35
barbara	156804	255	0	1
lena	153547	36	0	1

TABLE V
SELECTED TEST IMAGES WITH max. DECOMP. 7

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
artificial512	77937	20560	168	35
barbara	156820	271	0	1
lena	153562	51	0	1

size compared to the dyadic decomposition for RDO, the next column the further reduction for RDOE and the next column the further reduction of RDOH (compared to RDOE). For a decomposition depth of 5 compression efficiency is already improved for RDOH, RDOE, and RDO-WPB. Further increasing the decomposition depth reveals the difference between RDO and RDOH, the compression efficiency of RDO is decreased with increasing depth (as trend also followed by the dyadic decomposition) while RDOH can still remain its compression efficiency and even improve it. The improvement of compression efficiency is most pronounced for the Artificial and least significant for the Lena image.

Considering texture data, especially a well-performing subset of the Brodatz database, we see that highly textured data is well-suited for wavelet packet compression, almost 7KB are saved compared to the dyadic decomposition (see Tables VI–VIII). For lower decomposition depths RDO-WPB performs quite well, only at higher decomposition depths RDOE and RDOH can achieve performance improvements. The difference between RDOE and RDOH is small, thus the application of the proposed efficient header cost estimation (RDOE) is recommended.

B. Efficient Basis Selection

This section is dedicated to the analysis of more efficient wavelet packet basis selection than those based on actual rate-distortion optimization (RDOH, RDOE, and RDO-WPB). As

TABLE VI
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH max. DECOMP. 5

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
D101	295436	7117	0	1
D102	288898	13742	0	0
D103	345236	2871	0	0
D105	346221	6897	13	1
D106	356627	6891	22	19
D107	324110	6939	0	0
D109	333924	19822	0	0
D16	390561	14524	65	8
D21	362293	12675	271	51
D49	250951	14784	0	3
D53	319108	4987	96	18
D6	293454	5900	67	24
D64	275417	236	8	0
D67	317239	8655	0	1
D68	286124	1258	0	1
D76	327510	152	8	0
D77	365733	5391	19	28
D78	347068	2628	24	0
D79	334867	2093	16	0
D82	346751	41	0	5
D83	340033	2669	51	10
Avg.	326074.33	6679.61	31.42	8.09

discussed in Section V-B a top-down basis search is more efficient than a bottom-up best basis search, but may not lead to the best basis (optimal in the sense of an additive cost function). A top-down search with the Lagrangian rate has been conducted, the results are summarized in Figs. 15 and 16 in the graphs labelled “Top-down RDOH”. Most interestingly a top-down search leads to almost the same results as the more correct bottom-up solution for the FVC2004 database, while the performance is only slightly worse on the Brodatz database.

Entropy based basis selection works very well, but relies on the appropriate choice of the quantization parameter for a given image and rate, which is highlighted by the following results. The best performance of entropy based cost functions is summarized in the graphs with labels that contain “EX”; for this graph the best result of one of the quantization parameters 8, 16, 32, 48, and 64 has been chosen for every single image and rate (before averaging). These results can be compared to the performance of entropy based basis selection with a fixed quantization parameter ($q = 16$). The best results (EX) are very close to the actual optimum. E16 applied in bottom-up search performs very well on the Brodatz database, and at least significantly better than the dyadic basis on the FVC2004 database. The difference between a bottom-up search and a top-down search is almost negligible for EX (there are values of q for which a top-down search performs almost equally well than a

TABLE VII
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH max. DECOMP. 6

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
D101	295440	7121	0	1
D102	288904	13748	0	0
D103	345252	2887	0	0
D105	346225	6898	13	4
D106	356641	6894	33	19
D107	324135	6964	0	0
D109	333945	19843	0	0
D16	390574	14470	132	8
D21	362281	12263	747	139
D49	250956	14725	58	28
D53	319116	4764	294	54
D6	293437	5808	149	34
D64	275440	249	18	0
D67	317252	8668	0	1
D68	286137	1271	0	1
D76	327522	153	19	0
D77	365739	5370	40	34
D78	347078	2628	34	0
D79	334871	2093	16	4
D82	346772	62	0	5
D83	340030	2678	51	3
Avg.	326083.19	6645.57	76.38	15.95

TABLE VIII
SELECTED IMAGES FROM THE BRODATZ DATABASE WITH max. DECOMP. 7

Image	Dyadic	Reduc. RDO	F.r. RDOE	F.r. RDOH
D101	295454	7135	0	1
D102	288913	13757	0	0
D103	345271	2906	0	0
D105	346234	6898	22	4
D106	356654	6907	33	19
D107	324153	6982	0	0
D109	333959	19857	0	0
D16	390582	14457	153	8
D21	362292	12193	828	139
D49	250968	14670	125	28
D53	319127	4739	330	54
D6	293450	5798	172	34
D64	275456	265	18	0
D67	317268	8684	0	1
D68	286151	1285	0	1
D76	327536	167	19	0
D77	365754	5385	40	34
D78	347092	2642	34	0
D79	334883	2105	16	4
D82	346786	76	0	5
D83	340046	2683	62	3
Avg.	326096.61	6647.19	88.19	15.95

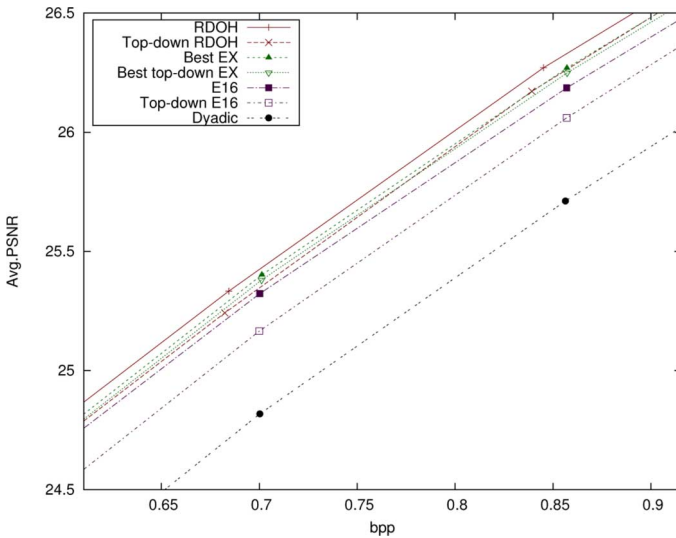


Fig. 15. Brodatz: Efficient basis selection.

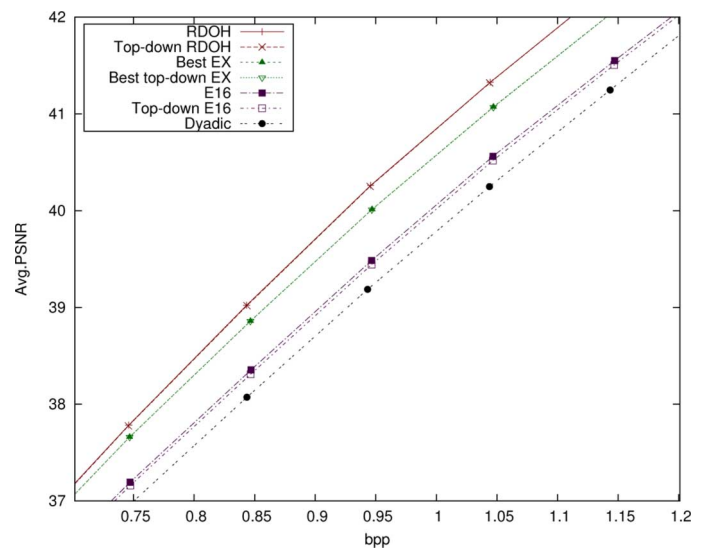


Fig. 16. FVC2004 DB4B: Efficient basis selection.

bottom up search), but for a fixed parameter (E16) the difference is more pronounced.

Entropy based basis selection performs significantly better than the standard dyadic wavelet packet basis and has been found to be superior to other cost functions such as the L2 norm (see Section V-A).

C. Anisotropic Wavelet Packet Bases

Anisotropic wavelet packet bases allow to adapt the WPB even closer to signal characteristics. However, as discussed in Section V-B, the complexity of rate-distortion optimal anisotropic WPB selection is tremendously higher than in the

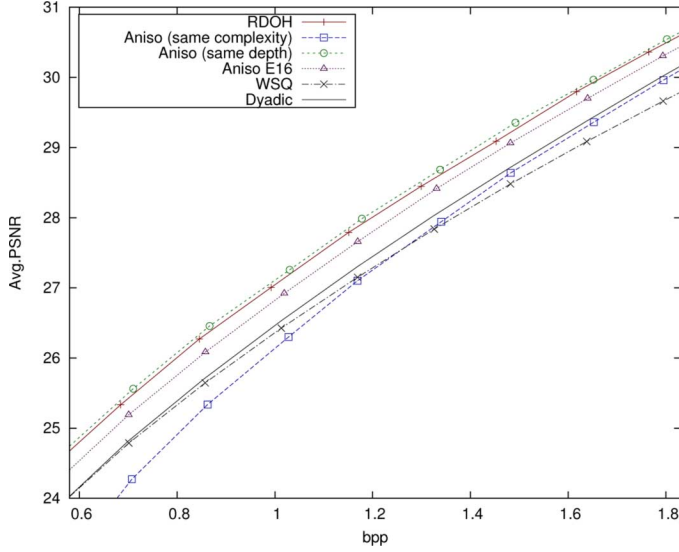


Fig. 17. Brodatz: PSNR evaluation.

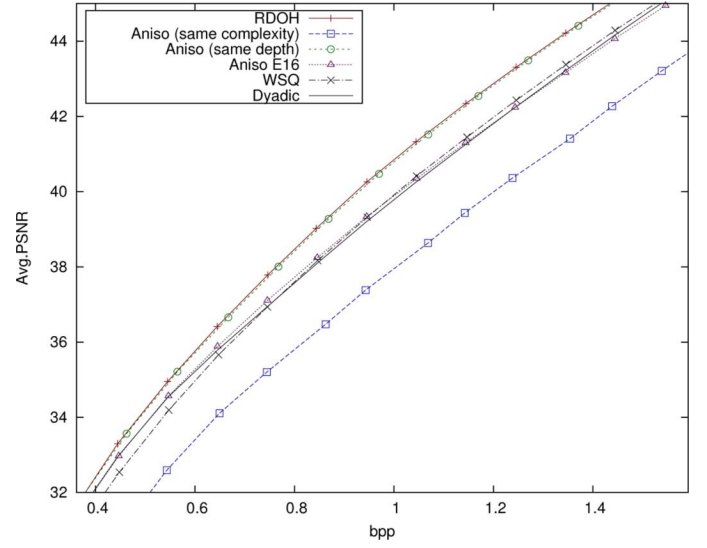


Fig. 19. FVC2004 DB4B: PSNR evaluation.

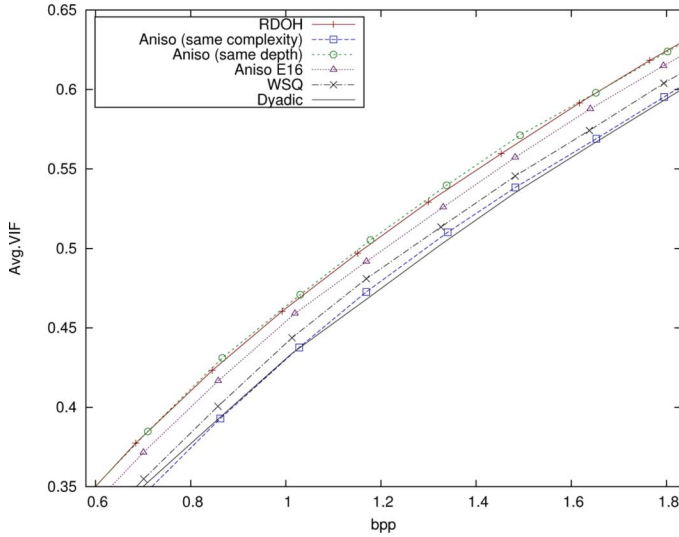


Fig. 18. Brodatz: VIF evaluation.

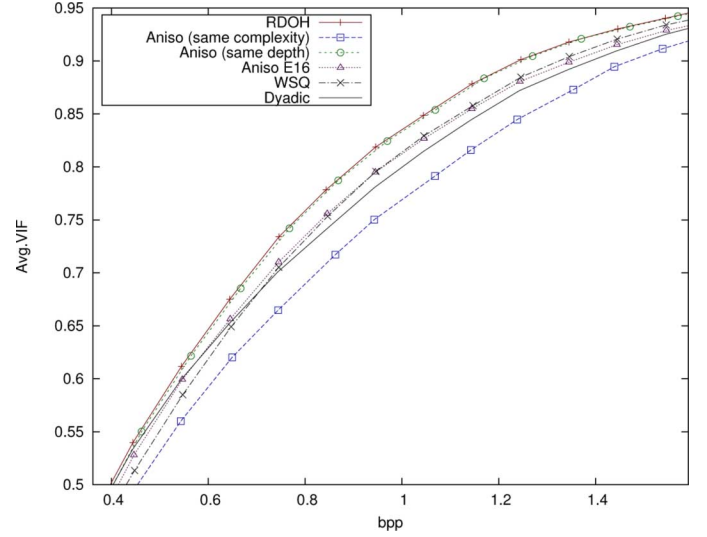


Fig. 20. FVC2004 DB4B: VIF evaluation.

isotropic case ($\mathcal{O}(d^2)$ compared to $\mathcal{O}(d)$). There are two ways to compare anisotropic WPB to isotropic WPB, based on the complexity to determine the basis or based on the complexity to compress the image with the given basis. If we follow the first approach (same selection complexity), we have to compare isotropic RDO WPB selection to anisotropic wavelet packet selection with a joint decomposition depth of 2, as in both cases the complexity is roughly equal (about six full JPEG2000 compressions).

If we choose the second approach, i.e., equal complexity in terms of compression with the computed optimal WPB, the equivalent anisotropic WPB has a maximum joint decomposition depth of 10 (for the isotropic decomposition depth 5). The complexity of anisotropic RDO WPB selection would roughly correspond to 66 full JPEG2000 compressions. Thus the complexity is too high for practical application and even straight forward computation. In order to estimate the performance of such deep decompositions we give the best performance of all

results, i.e., from cost functions in the anisotropic case and the isotropic RDOH. Given the results of the best performance of cost functions in the isotropic case this approach seems well justified (see Figs. 15 and 16).

At the same complexity anisotropic WPB selection does not lead to competitive results at all (see Figs. 17 and 19), while at the same depth slight performance improvements are achievable. However, the increased complexity does not justify the achieved performance.

Additionally to our best WPB selection algorithms for JPEG2000 we evaluated the compression performance of the WSQ WPB, which led to even worse results than the standard dyadic decomposition on the Brodatz database (all bitrates) and for lower bitrates on the FVC2004 database. It is notable that the VIF quality metric reports higher quality scores for the WSQ WPB (see Figs. 18 and 20), which justifies the application of the WSQ from a more human perception oriented point of view.

Our results have also revealed that even when the restrictions of permissible WPB in JPEG2000 Part 2 are applied the excellent compression performance of rate-distortion optimal wavelet packet basis selection is preserved.

VII. CONCLUSION

Rate-distortion optimal (RDO) wavelet packet basis (WPB) selection for JPEG2000 has been presented and discussed.

Our algorithms enable the selection of the best WPB in the JPEG2000 coding framework. Thus we are able to report the upper bound of performance improvements achievable with custom wavelet packet bases in JPEG2000 and the performance impairment of more efficient heuristics. We could also show that the restrictions on the set of permissible WPB as standardized in JPEG2000 Part 2 preserve the excellent compression performance of rate-distortion optimal wavelet packet selection.

In terms of compression performance our results show that for highly textured data, RDO wavelet packet bases perform significantly better than the dyadic decomposition. Compression performance improvements can be reported even for the lossless case. For higher decomposition depths the consideration of the header cost in JPEG2000 WPB optimization is favorable. We have presented an efficient algorithm for packet header estimation (RDOE) that performs almost equal to actual header cost computation (by simulating compression). The improved PSNR performance of RDO WPB also leads to improvements in terms of state-of-the-art objective image quality assessment. For highly textured image data significant compression performance improvements are shown.

The isotropic RDOE algorithm offers a computationally reasonable way to improve compression performance. It takes the header data into account and thus efficiently prevents inferior compression results at greater decomposition depths.

REFERENCES

- [1] R. R. Coifman and M. V. Wickerhauser, "Entropy based methods for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 719–746, 1992.
- [2] C. Taswell, "Satisficing search algorithms for selecting near-best bases in adaptive tree-structured wavelet transforms," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2423–2438, Oct. 1996.
- [3] T. Schell and A. Uhl, "Optimization and assessment of wavelet packet decompositions with evolutionary computation," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 8, pp. 806–813, 2003.
- [4] K. Ramchandran and M. Vetterli, "Best wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 160–175, Apr. 1993.
- [5] N. M. Rajpoot, R. G. Wilson, F. G. Meyer, and R. R. Coifman, "Adaptive wavelet packet basis selection for zerotree image coding," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1460–1472, Dec. 2003.
- [6] Y. Yang and C. Xu, "A wavelet packet based block-partitioning image coding algorithm with rate-distortion optimization," in *Proc. IEEE Int. Conf. Image Processing (ICIP'05)*, 2005, vol. III, pp. 201–204.
- [7] Y. Yang and C. Xu, "Fast and efficient basis selection methods for embedded wavelet packet image coding," in *Proc. 3rd Int. Conf. Image Analysis and Recognition (ICIAR'06)*, Sep. 2006, vol. 4141, Lecture Notes in Computer Science, pp. 480–492.
- [8] Y. Yang and C. Xu, "Fast wavelet packet basis selection for block-partitioning image coding," in *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2006)*, Sep. 2006, 4 pp.
- [9] J. Yang, Y. Wang, W. Xu, and Q. Dai, "Image coding using dual-tree discrete wavelet transform," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1555–1569, Sep. 2008.
- [10] J. N. Bradley, C. M. Brislawn, and T. Hopper, "The FBI wavelet/scalar quantization standard for gray-scale fingerprint image compression," in *Proc. SPIE Visual Information Processing II*, Orlando, FL, Apr. 1993, vol. 1961, pp. 293–304.
- [11] R. C. Kidd, "Comparison of wavelet scalar quantization and JPEG for fingerprint image compression," *J. Electron. Imag.*, vol. 4, no. 1, pp. 31–39, 1995.
- [12] S. Kasaei, M. Deriche, and B. Boashash, "A novel fingerprint image compression technique using wavelet packets and pyramid lattice vector quantization," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1365–1378, Nov. 2002.
- [13] P. Saeedian and B. Shirazi, "A novel fingerprint image compression technique using adaptive subband image coding," in *Proc. 2004 Picture Coding Symp. (PCS'04)*, 2004.
- [14] G. A. Khuwaja, "Best parameter based compression of fingerprints with wavelet packets," *Int. J. Comput. Appl. Technol.*, vol. 19, no. 1, pp. 51–62, 2004.
- [15] D. Taubman and M. W. Marcellin, *JPEG2000—Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer, 2002.
- [16] *Information Technology—JPEG2000 Image Coding System, Part 2: Extensions*, ISO/IEC 15444-2, 2004.
- [17] M. Reisecker and A. Uhl, "Wavelet-packet subband structures in the evolution of the JPEG2000 standard," in *Proc. 6th IEEE Nordic Signal Processing Symp. (NORSIG 2004)*, CD-ROM, Espoo, Finland, Jun. 2004.
- [18] D. Engel and A. Uhl, "Secret wavelet packet decompositions for JPEG2000 lightweight encryption," in *Proc. 31st Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '06)*, Toulouse, France, May 2006, vol. V, pp. 465–468.
- [19] T. Stütz, B. Mühlbacher, and A. Uhl, "Best wavelet packet bases in a JPEG2000 rate-distortion sense: The impact of header data," in *Proc. IEEE Int. Conf. Multimedia & Expo (ICME '10)*, Singapore, Jul. 2010, pp. 19–24.
- [20] B. Mühlbacher, T. Stütz, and A. Uhl, P. Frossard, H. Li, F. Wu, B. Girod, S. Li, and G. Wei, Eds., "JPEG2000 part 2 wavelet packet subband structures in fingerprint recognition," in *Proc. SPIE Visual Communications and Image Processing 2010 (VCIP'10)*, Huang Shan, China, Jul. 2010, pp. 77442C-1–77442C-10.
- [21] *Information Technology—JPEG2000 Image Coding System, Part 2: Extensions*, ITU-T T.801, 2002.
- [22] D. Xu and M. N. Do, "Anisotropic 2-D wavelet packets and rectangular tiling: Theory and algorithms," in *Proc. SPIE Conf. Wavelets X*, 2003, vol. 19, pp. 619–630.
- [23] *Information Technology—JPEG2000 Image Coding System, Part 1: Core Coding System*, ISO/IEC 15444-1, 2000.
- [24] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1992.
- [25] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.



Thomas Stütz received the Ph.D. degree in applied computer science. His Ph.D. thesis focused on format-specific multimedia encryption, i.e., special encryption algorithms that preserve certain properties of the multimedia format, such as scalability.

He is a lecturer at the Salzburg University of Applied Sciences, Salzburg, Austria, and currently has a PostDoc position at the Polytech' Nantes, IRCCyN-IVC, Nantes, France. His main research interests are compression and multimedia security.



Andreas Uhl is an Associate Professor at the Computer Sciences Department of the University of Salzburg, Salzburg, Austria, where he leads the Multimedia Signal Processing and Security Lab. He is also lecturer at the Carinthia Tech Institute and the Salzburg University of Applied Sciences. His research interests include image and video processing, wavelets, multimedia security, biometrics, parallel algorithms, and number-theoretical numerics.